

LAPORAN PENELITIAN

PENYETARAAN PERANGKAT TES MATEMATIKA PROGRAM D2 PGSD UNIVERSITAS TERBUKA

Oleh:

Rustam

LEMBAGA PENELITIAN UNIVERSITAS TERBUKA

2000



LEMBAR PENGESAHAN
LAPORAN PENELITIAN LEMBAGA PENELITIAN-UT

- | | |
|---------------------------|---|
| 1. a. Judul Penelitian | : Penyetaraan Perangkat Tes Matematika Program D2 PGSD Universitas Terbuka |
| b. Bidang Penelitian | : Penelitian Keilmuan |
| c. Klasifikasi Penelitian | : Penelitian Mandiri. |
| d. Bidang Ilmu | : Evaluasi/Pengukuran Pendidikan. |
| 2. Ketua Peneliti | |
| a. Nama Lengkap dan gelar | : Drs. Rustam, M.Pd. |
| b. NIP. | : 131 925 717 |
| c. Golongan kepangkatan | : III/d. |
| d. Jabatan akademik | : Lektor Madya. |
| e. Fakultas/Unit Kerja | : FKIP |
| 3. Lama Penelitian | : Delapan bulan. |
| 4. Biaya Penelitian | : Rp. 3.817.500,00 (Tiga juta delapan ratus tujuh belas ribu lima ratus rupiah). |
| 5. Sumber Biaya | : Universitas Terbuka. |

Mengetahui,

Dekan/FKIP



Drs. H. Udin S. Winataputra, M.A.

NIP. 130 367 151

Mengetahui,

Ketua Lembaga Penelitian UT



WBP. Simanjuntak, M.Ed., Ph.D

NIP. 130 212 017

Pondok Cabe, 26 Desember 2000.

Ketua Peneliti,

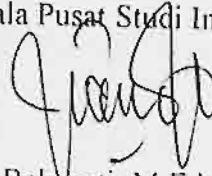


Drs. Rustam, M.Pd.

NIP. 130 925 717

Menyetujui,

Kepala Pusat Studi Indonesia



Tian Belawati, M.Ed Ph.D

NIP. 131 569 974.

ABSTRAK

Rustam (2001). *Penyetaraan Perangkat Tes Matematika Program D2 PGSD Universitas Terbuka*

Laporan penelitian. Jakarta: Lembaga Penelitian Universitas Terbuka.

Penelitian ini bertujuan untuk menemukan persamaan konversi skor (formula equating) hasil penyetaraan dari dua perangkat tes matakuliah PGSD 2204 yang disusun berdasarkan kisi-kisi yang sama.

Populasi dalam penelitian ini adalah mahasiswa Program D2 PGSD yang mengambil matakuliah Matematika (PGSD 2204). Sampel sebanyak 139 mahasiswa yang diambil dari dua UPBJJ, yaitu UPBJJ Bogor dan UPBJJ Bandar Lampung. Pengambilan sampel menggunakan teknik *Purposive Cluster Sampling*.

Hasil analisis data menunjukkan bahwa persamaan konversi skor (formula equating) dari dua perangkat tes matakuliah Matematika (PGSD 2204), yaitu antara perangkat tes P dan perangkat tes Q adalah $Q^* = 0,909P + 6,136$ atau $P^* = 1,100Q - 6,748$, dengan masing-masing kesalahan pengukuran penyetaraan $SEE[Q(P_0)] = 1,151$ dan $SEE[P(Q_0)] = 1,263$. Berdasarkan kedua persamaan konversi dan kesalahan pengukuran penyetaraan di atas dapat disimpulkan bahwa persamaan konversi yang terbaik untuk digunakan adalah $Q^* = 0,909P + 6,136$; dan perangkat tes P lebih sukar dibandingkan dengan perangkat tes Q.

DAFTAR ISI

Lembar Pengesahan	i
Abstrak	ii
Daftar Isi	iii
Daftar Tabel	iv
BAB I : Pendahuluan	
A. Latar Belakang	1
B. Batasan dan Rumusan Masalah	4
C. Tujuan Penelitian	4
D. Manfaat Penelitian	5
BAB II : Tinjauan Pustaka	
A. Karakteristik Tes	6
B. Penyetaraan Perangkat Tes	8
BAB III : Metodologi Penelitian	
A. Tempat dan Waktu Penelitian	18
B. Instrumen Penelitian	18
C. Populasi dan Sampel Penelitian	19
D. Metode Pengumpulan Data	19
E. Metode Analisis Data	20
BAB IV : Hasil Penelitian dan Pembahasan	
A. Hasil Penelitian	21
B. Pembahasan	28
BAB V : Kesimpulan dan Saran	
A. Kesimpulan	32
B. Saran	32
Daftar Pustaka	
Lampiran : 1. Data Respons Peserta Tes Kelompok I (K_1)	
2. Data Respons Peserta Tes Kelompok II (K_2)	
3. Skor Peserta Tes Kelompok I (K_1)	
4. Skor Peserta Tes Kelompok II (K_2)	
5. Identitas Peneliti.	

DAFTAR TABEL

Tabel 1. Skema Rancangan Penyetaraan Tes.

Tabel 2. Nilai Statistik yang Dibutuhkan pada Penyetaraan Perangkat Tes Rancangan C.

Tabel 3. Karakteristik Tes Perangkat Tes P,Q, dan R.

Tabel 4. Data Nilai Statistik Skor Peserta Tes pada Perangkat Tes P,Q, dan R.

Tabel 5. Nilai Statistik untuk Menentukan Kesalahan Pengukuran Penyetaraan Perangkat Tes P dan Q.

Tabel 6. Kesalahan Pengukuran Penyetaraan dari Perangkat Tes P ke Perangkat Tes Q.

Tabel 7. Kesalahan Pengukuran Penyetaraan dari Perangkat Tes Q ke Perangkat Tes P.

BAB I

PENDAHULUAN

A. Latar Belakang Masalah

Penilaian di dalam kegiatan belajar mengajar dimaksudkan untuk mendapatkan informasi tentang sejauh mana keberhasilan guru dalam menyajikan bahan pelajaran, serta sejauh mana peserta didik dapat menyerap materi yang diajarkan. Objektivitas penilaian hasil belajar yang dilakukan sangat bergantung pada kualitas dari alat ukur (tes) yang digunakan. Umar (1991) menyatakan, bahwa hanya penilaian yang didasarkan kepada hasil pengukuran yang dapat dipercaya sajalah yang dapat dijadikan sebagai landasan yang kuat bagi pengambil keputusan tentang keberhasilan pengajar maupun peserta didik. Selanjutnya, Thorndike & Hagen (1977) menegaskan, bahwa keputusan akan lebih sempurna apabila untuk mendapatkan informasi tersebut digunakan alat ukur yang baik, sehingga informasi yang diperoleh betul-betul objektif dan tidak bias.

Kenyataan menunjukkan bahwa sering terjadi pemberian nilai terhadap hasil belajar siswa tidak mencerminkan prestasi sebenarnya. Hal ini terutama disebabkan oleh tes yang digunakan dalam penilaian tidak mencerminkan alat ukur yang benar-benar dapat dipercaya. Alat ukur yang digunakan kemungkinan tidak cocok, sehingga nilai yang diberikan kepada siswa atau peserta tes lebih tinggi atau lebih rendah dari prestasi sebenarnya. Hal lain yang sering terjadi adalah dalam menafsirkan hasil pengukuran, dua atau beberapa perangkat tes sering diperlakukan sama, tanpa memperhatikan tingkat kesukaran perangkat tes yang digunakan.

Guru atau dosen dalam menyusun dan melaksanakan tes untuk kepentingan penilaian hasil belajar cenderung lemah dalam menafsirkan hasil pengukuran dari dua atau beberapa perangkat tes yang berbeda tetapi mengukur hal yang sama, mereka sering memperlakukan hal tersebut sama, tanpa memperhatikan tingkat kesukaran perangkat tes yang digunakan. Menurut Hamblethor & Swaminathan (1985) tindakan

demikian dapat dikatakan menyimpang dari teori pengukuran, karena sebenarnya tidak pernah ditemukan dua perangkat tes yang mempunyai butir soal berbeda, sama tingkat kesukarannya.

Menurut Suryabrata (1987) bahwa dalam pelaksanaan evaluasi yang menggunakan beberapa perangkat tes yang berbeda dan mengukur hal yang sama perlu dilakukan penyetaraan dari beberapa perangkat tes tersebut, karena dengan adanya penyetaraan dapat dijamin keadilan bagi peserta tes. Selanjutnya, Hamblethton & Swaminathan (1985) mengingatkan bahwa sekalipun perangkat tes yang disusun berdasarkan kisi-kisi yang sama, jarang sekali bahkan hampir tidak ditemukan perangkat tes yang benar-benar setara dalam sebaran serta peringkat indeks kesukaran. Dengan demikian, menjadi suatu keharusan bagi para pengembang tes atau lembaga pengembang tes untuk melakukan penyetaraan terhadap perangkat tes yang digunakan.

Kebutuhan terhadap pentingnya penyetaraan tes telah lama disadari, namun pemikiran dan penelitian terhadap penyetaraan tes masih sangat terbatas. Beberapa penelitian yang dilakukan menemukan persamaan konversi dari beberapa perangkat tes yang telah diterapkan. Surapranata (1993) hasil penelitiannya, ditemukan suatu formulasi equiting dari dua perangkat tes ebtanas SMP tahun ajaran 1991/1992 di DKI Jakarta yang digunakan pada subrayon yang berbeda. Selanjutnya, Snulling (1993) hasil penelitiannya menunjukkan, bahwa semua perangkat tes ebtanas SD yang digunakan di Sumatra Selatan pada tahun 1991/1992 tidak paralel dibandingkan dengan dengan perangkat tes ebtanas SD yang digunakan di Kotamadya Palembang. Kedua hasil penelitian tersebut menunjukkan bahwa skor pada perangkat tes yang berbeda tidak dapat disamakan atau dibandingkan langsung, tetapi harus melalui proses penyetaraan terhadap perangkat tes yang digunakan.

Melalui penyetaraan diperoleh dua keuntungan pokok, yaitu (1) dimungkinkannya penggunaan perangkat tes yang berbeda terhadap kelompok yang berbeda sesuai dengan tingkat kemampuannya, sehingga skor yang diperoleh dapat dibandingkan dan peserta tes tidak merasa dirugikan atau diuntungkan karena mereka

kebetulan mengambil perangkat tes yang lebih sukar atau lebih mudah; (2) bila terjadi kebocoran soal dari suatu perangkat tes tertentu dapat segera diganti dengan perangkat tes lain yang telah tersedia.

Penyetaraan tes sangat dirasakan kegunaannya mengingat mutu pendidikan belum merata, dan keadaan geografis wilayah Indonesia sebagai negara kepulauan yang cukup luas. Mutu pendidikan di daerah pedesaan berbeda dengan mutu pendidikan di daerah perkotaan, demikian pula mutu pendidikan di pulau Jawa cenderung berbeda dengan mutu pendidikan di luar pulau Jawa, dan tidak dapat disamakan. Hal ini dikarenakan sekolah di perkotaan atau di pulau Jawa didukung oleh fasilitas yang cukup memadai, sedangkan di pedesaan atau di luar pulau Jawa pendukungnya relatif kurang. Dalam kasus demikian, tidaklah bijaksana bila untuk mengukur kemampuan peserta tes pada daerah atau sekolah yang mutunya sangat berbeda diberikan perangkat tes yang sama. Kita akan lebih adil bila prestasi hasil belajar peserta tes pada daerah atau sekolah yang berbeda diukur dengan perangkat tes yang berbeda pula, dengan mempertimbangkan tingkat kemampuan mereka.

Universitas Terbuka (UT) merupakan satu-satunya lembaga Pendidikan Tinggi di Indonesia yang mengelola ujian akhir semester (UAS) secara terpusat. Matakuliah yang ditawarkan setiap semester sebanyak 900 matakuliah. Sementara mahasiswa aktif sampai Agustus 2000 sebanyak 314.698 orang (Depdiknas, 2000). Melihat kondisi ini sangatlah mungkin kualitas mahasiswa berbeda karena mereka kebetulan tinggal daerah terpencil dan juga sangatlah mungkin perangkat tes UAS 'bocor' sebelum digunakan. Oleh karena itu, UT telah saatnya perlu mengembangkan perangkat tes setara.

Uraian di atas memberi gambaran bahwa penyetaraan perangkat tes sebagai salah satu upaya pengembangan alat ukur dalam mengevaluasi hasil belajar perlu dikembangkan dalam berbagai bidang ilmu. Khusus di UT, matakuliah yang perlu mendapatkan prioritas utama dalam pengembangan perangkat tes yang setara adalah matakuliah yang banyak direspons mahasiswa dan yang dianggap sukar bagi kebanyakan mahasiswa.

B. Batasan dan Rumusan Masalah

Sasaran utama penelitian ini adalah penyetaraan dua perangkat tes yang disusun berdasarkan kisi-kisi yang sama. Hasil pelaksanaan penyetaraan diperoleh dua perangkat tes yang memiliki nilai kesetaraan tertentu, sehingga dapat digunakan sebagai alat ukur yang baik dan adil dalam penilaian hasil belajar peserta tes.

Penyetaraan perangkat tes, secara garis besar dibedakan antara metode penyetaraan horizontal dan metode penyetaraan vertikal (Naga, 1992). Penyetaraan horizontal dibagi atas dua macam, yaitu metode linear dan metode ekuipersentil. Mengingat luasnya teori yang terkait dengan penyetaraan tes, pada penelitian ini hanya dibatasi pada penyetaraan horizontal dengan menggunakan metode linear. Penyetaraan horizontal yang dimaksud adalah menyetarakan dua (kelompok) skor yang masing-masing diperoleh dari dua perangkat tes yang berbeda (perangkat tes tipe I dan II) akan tetapi mengukur hal yang sama, yakni kemampuan peserta tes (mahasiswa) pada matakuliah Matematika (PGSD 2204). Sedangkan pendekatan linear dimaksudkan proses penyetaraan skor beberapa perangkat tes berhubungan dengan fungsi dan regresi linear.

Berdasarkan uraian di atas didapat rumusan masalah sebagai berikut. Bagaimanakah persamaan konversi atau formula penyetaraan hasil penyetaraan dua perangkat tes matakuliah PGSD 2204 yang disusun berdasarkan kisi-kisi yang sama?

C. Tujuan Penelitian

Tujuan penelitian ini adalah untuk menemukan formula atau rumus konversi skor dari dua perangkat tes matakuliah PGSD 2204 yang disusun berdasarkan kisi-kisi yang sama.

D. Manfaat Penelitian

Hasil penelitian ini diharapkan menambah khasanah empirik dalam rangka pengembangan tes hasil belajar, terutama yang terkait dengan penyetaraan perangkat tes. Selain itu secara khusus, temuan penelitian ini dapat menjadi bahan masukan bagi pihak-pihak yang terlibat dalam pengembangan tes, terutama yang terkait dengan penyelenggaraan UAS di Universitas Terbuka.

BAB II

TINJAUAN PUSTAKA

A. Karakteristik Tes

Karakteristik tes untuk tes hasil belajar dapat dibedakan menjadi dua kelompok kajian, pertama karakteristik butir tes dan karakteristik perangkat tes yang merupakan kumpulan dari beberapa butir tes. Karakteristik butir tes yang pada umumnya digunakan untuk menentukan butir itu baik atau tidak biasanya menggunakan dua bentuk parameter, yaitu parameter tingkat kesukaran dan daya beda butir soal. Sedangkan karakteristik perangkat tes umumnya yang digunakan sebagai kriteria baik tidaknya (perangkat) tes menggunakan parameter tingkat kesukaran tes, daya beda tes, reliabilitas, dan kesalahan pengukuran standar atau *Standard Error of Measurement (SEM)*.

1. Karakteristik butir soal

a. Tingkat kesukaran butir soal

Tingkat kesukaran butir soal didefinisikan sebagai proporsi subyek yang menjawab butir tes tertentu dengan benar. Sedangkan angka yang menunjukkan sukar mudahnya suatu butir soal disebut indeks kesukaran (p) dengan nilai antara 0 dan 1. Nilai 0 berarti semua peserta tes tidak dapat menjawab butir tersebut dengan benar (butir soal tersebut sangat sukar). Sedangkan nilai 1 berarti semua peserta tes dapat menjawab butir tersebut dengan benar (butir soal tersebut sangat mudah). Butir soal yang terlalu sukar atau terlalu mudah sebaiknya dihindari karena butir soal tersebut kurang mencerminkan alat ukur yang baik.

b. Daya beda

Daya beda butir soal merupakan kemampuan suatu butir soal untuk membedakan antara peserta tes yang pandai dengan peserta tes yang kurang pandai.

Suryabrata (1999) menyatakan, tujuan pokok menentukan daya beda butir soal adalah untuk menentukan apakah butir soal tersebut memiliki kemampuan membedakan kelompok dalam aspek yang diukur, sesuai dengan perbedaan yang ada pada kelompok tersebut. Daya beda butir soal yang sering digunakan dalam tes hasil belajar adalah dengan cara menggunakan indeks korelasi antara skor butir dengan skor totalnya. Daya beda dengan cara ini sering disebut validitas internal, karena korelasi diperoleh dari dalam tes itu sendiri. Teknik korelasi yang sering digunakan untuk menghitung daya beda adalah teknik point biserial dan teknik biserial.

Koefisien daya beda bergerak dari -1 sampai $+1$. Apabila suatu butir memiliki korelasi negatif dapat dikatakan butir tersebut menyesatkan, karena subyek yang terdiri dari kelompok pandai lebih menjawab salah dari pada subyek yang termasuk kelompok kurang pandai, atau kelompok pandai lebih sedikit menjawab benar dibandingkan dengan kelompok subyek yang kurang pandai. Menurut Fernandes (1984) mengemukakan bahwa butir soal yang mempunyai indeks daya beda $0,20$ atau lebih besar, butir soal tersebut sudah dapat dikatakan baik.

2. Karakteristik (perangkat) tes

a. Tingkat kesukaran dan daya beda tes

Tingkat kesukaran dan daya beda selain sebagai karakteristik butir soal, juga dapat menjadi karakteristik (perangkat) tes, yaitu dengan cara merata-ratakan seluruh indeks kesukaran dan indeks daya beda butir soal yang membangun tes tersebut. Dengan kata lain, Tingkat kesukaran dan daya butir tes dapat dilihat dari rerata indeks kesukaran dan rerata daya beda butir-butir soal yang membentuk tes tersebut.

b. Reliabilitas

Selain tingkat kesukaran dan daya beda tes, kualitas tes secara keseluruhan dapat dilihat pada besarnya koefisien reliabilitas perangkat tes. Reliabilitas dapat diartikan dengan tingkat kepercayaan dari suatu alat ukur. Artinya, sampai seberapa jauh pengukuran yang dilakukan berulang-ulang terhadap subyek atau sekelompok

subyek yang sama memberikan hasil yang relatif sama (tetap). Ukuran besarnya reliabilitas menurut Gronlund (1982) telah cukup memadai jika besarnya sekitar 0,60 sampai dengan 0,85.

c. Kesalahan pengukuran standar (Standard Error of Measurement-SEM)

Unsur lain yang menjadi parameter baik tidaknya suatu tes adalah kesalahan pengukuran standar (SEM). Kesalahan pengukuran dalam tes disebabkan oleh kesalahan pengambilan sampel (sampling error) peserta tes dan kesalahan pelaksanaan tes itu sendiri (measurement error). Besarnya nilai kesalahan pengukuran suatu perangkat tes, Hopkins, dkk. (1990) mengemukakan rumus :

$$SEM = \sigma \sqrt{1 - r_{xx}}$$

Keterangan:

σ = simpangan baku skor tes

r_{xx} = koefisien reliabilitas tes.

Nilai *SEM*, semakin kecil semakin baik. Artinya, skor yang diperoleh dari perangkat tes tersebut semakin mendekati skor yang sebenarnya.

B. Penyetaraan Perangkat Tes

1. Pengertian penyetaraan

Menurut Angoff (dalam Holland & Rubin, 1982) penyetaraan merupakan proses konversi dari sistem unit tes yang satu ke sistem unit tes yang lain, agar skor-skor yang berasal dari dua perangkat tes setelah dikonversi adalah sama. Crocker & Algina (1986) mengatakan bahwa penyetaraan adalah dua skor, yang satu perangkat tes X dan lainnya perangkat tes Y dianggap ekuivalen, jika perangkat tes X dan perangkat tes Y mengukur ciri yang sama dengan reliabilitas relatif sama, serta peringkat persentil yang sepadan dengan sebaran skor. Selanjutnya, Petersen, Marco, dan Stewart (dalam Holland & Rubin, 1982) menyatakan, penyetaraan adalah suatu prosedur empiris yang diperlukan untuk mentransformasi skor dari tes yang satu ke

skor tes yang lain. Dengan demikian, penyetaraan adalah suatu cara untuk mengubah skor-skor dari perangkat tes yang berbeda dalam rerata tingkat kesukaran dan rerata daya beda butir di antara dua atau lebih perangkat tes, agar menjadi sama setelah dilakukannya penyetaraan.

Berdasarkan beberapa pengertian yang telah dikemukakan di atas, dapat dikatakan bahwa penyetaraan atau *equating* merupakan suatu prosedur secara empiris dalam rangka menyetarakan skor dari perangkat tes yang satu ke perangkat tes yang lain. Melalui proses penyetaraan yang dilakukan dengan benar, dapat diukur perkembangan dan memetakan berbagai kecenderungan, serta mendapatkan dan menggabungkan informasi dari perangkat tes yang berbeda. Hasil penyetaraan memungkinkan kita melakukan perbandingan atau konversi secara langsung hasil-hasil tes individu yang mengikuti perangkat tes yang berbeda.

2. Syarat-syarat penyetaraan

Syarat utama agar supaya skor dari dua perangkat tes dapat disetarakan adalah kedua perangkat tes tersebut berasal dari kisi-kisi yang sama. Persyaratan lain yang juga harus dipenuhi adalah kedua perangkat tes yang disetarakan memiliki reliabilitas sama. Namun demikian, Angoff (dalam Holland dan Rubin, 1982) menyatakan, bahwa skor dari dua perangkat tes yang memiliki reliabilitas tidak sama juga dapat disetarakan. Menurut Petersen *et al.* (dalam Holland & Rubin, 1982), skor-skor dari perangkat tes X dan Y dapat disetarakan jika terpenuhi empat syarat, yaitu (1) kedua tes harus mengukur hal yang sama, (2) kemampuan peserta tes harus sama, (3) transformasinya harus invarian, dan (4) transformasinya harus simetris.

Persyaratan invariansi dan simetris bertitik tolak pada tujuan penyetaraan, yaitu untuk menghasilkan skor-skor yang ekuivalen, yang berasal dari dua atau lebih perangkat tes. Jika skor-skor pada dua perangkat tes itu ekuivalen, maka adanya keadilan antara dua perangkat tes tersebut. Dengan demikian, konversi skor tes memerlukan transformasi yang dapat dibalik, yaitu jika skor Y_0 pada perangkat tes Y

disetarakan dengan skor X_0 pada perangkat tes X, maka skor X_0 pada perangkat tes X harus dapat disetarakan dengan skor Y_0 pada perangkat tes Y.

3. Macam-macam penyetaraan

Ditinjau dari peserta tes, penyetaraan dibedakan atas dua macam, yaitu penyetaraan horizontal dan penyetaraan vertikal (Naga, 1992). Penyetaraan horizontal dimaksudkan untuk menyetarakan dua skor yang masing-masing diperoleh dari dua perangkat tes yang berbeda, tetapi mengukur hal yang sama. Penyetaraan vertikal dimaksudkan untuk menentukan padanan skor-skor yang diperoleh dari dua kelompok peserta tes dalam tingkat atau jenjang pendidikan yang berbeda, akan tetapi dikenakan perangkat tes yang sama.

Penyetaraan vertikal dilakukan untuk mengetahui perkembangan kemampuan peserta tes. Misalnya, ditemukan peserta tes dari kelas yang lebih rendah, karena suatu hal peserta tersebut tidak cukup mampu (terbelakang) dari teman sekelasnya, sehingga tidak cukup mampu mengikuti ujian (tes) dari kelasnya sendiri. Jika dipaksakan ikut tes sekelasnya, maka peserta itu tidak dapat menjawab dengan benar hampir semua butir soal dalam perangkat tes yang disediakan, sehingga hasil ukur dari perangkat tes itu akan memiliki informasi yang sangat rendah. Untuk memperoleh informasi yang cukup, kepada peserta tes yang terbelakang tersebut diberikan perangkat tes dari kelas yang lebih rendah. Kemudian skor hasil perangkat tes tersebut perlu dikembalikan ke skor kelasnya.

4. Prosedur Penyetaraan

Penyetaraan dilakukan dengan cara mengkonversikan sistem unit suatu perangkat ke sistem unit perangkat tes lain, dari tes yang mengukur kemampuan yang sama. Penyetaraan perangkat tes merupakan pembuatan sejumlah keputusan dari skor yang diperoleh pada sebuah perangkat tes untuk disesuaikan ke bentuk yang berbeda dalam tingkat kesukarannya. Berarti, nilai yang didapat dari tes yang lebih sulit akan lebih berharga dari nilai yang didapat dari tes yang lebih mudah (Suryabrata, 1987).

Menurut Petersen *et al.* (dalam Holland & Rubin, 1982), masalah penyetaraan tidak sederhana seperti masalah regresi, karena metode penyetaraan adalah sebuah prosedur empiris yang melibatkan sebuah desain untuk pengumpulan data dan sebuah aturan untuk menetapkan transformasi. Naga (1992) mengemukakan tiga desain dasar untuk mengumpulkan data, yaitu (1) metode kelompok tunggal, (2) metode kelompok ekuivalen, dan (3) metode anchor tes; dan terdapat tiga aturan untuk menetapkan transformasi, yaitu (1) penyetaraan ekipersentil, (2) penyetaraan linear, dan (3) kurva karakteristik soal. Selanjutnya, Hamblethton, Swaminathan, & Roger (1991) mengemukakan bahwa, pada teori tes klasik dikenal dua macam metode penyetaraan, yakni metode penyetaraan linear, dan metode penyetaraan ekipersentil. Metode penyetaraan linear akan menghubungkan skor konversi dengan skor asalnya melalui suatu fungsi linear. Sedangkan metode penyetaraan ekipersentil, mengkonversi skor berdasarkan jenjang persentil yang sama.

Petersen *at al.* (dalam Holland & Rubin, 1982) mengemukakan tiga jenis rancangan dalam penyetaraan skor di antara beberapa perangkat tes yang berbeda, yaitu rancangan A, rancangan B, dan rancangan C. Dalam penelitian ini, dibatasi pada metode yang paling sederhana, yakni hanya menggunakan dua perangkat tes dan dua kelompok peserta tes. Skema rancangan, perangkat tes yang digunakan, dan kelompok yang dilibatkan seperti berikut.

Tabel 1. Skema Rancangan Penyetaraan Tes.

Jenis Rancangan	Kelompok Peserta Tes	
	K ₁	K ₂
A	P	Q
B	P,Q	Q,P
C	P+R	Q+R

Keterangan: Diadaptasi dari Crocker & Algina, 1986.

Pada rancangan A, peserta dibagi secara acak ke dalam dua kelompok, yaitu kelompok pertama (K_1) dan kelompok kedua (K_2). Kelompok acak dimaksudkan agar peserta tes K_1 dan K_2 adalah setara yakni memiliki kemampuan yang sama dengan variansi yang sama pula. Setelah itu K_1 mengerjakan perangkat tes P dan K_2 mengerjakan perangkat tes Q (lihat Tabel 1).

Pada rancangan B, peserta tes dibagi dalam K_1 dan K_2 dan diperoleh secara acak. Cara acak ini dimaksudkan agar kedua kelompok setara yakni mereka memiliki kemampuan yang sama dengan variansi yang sama pula. Setelah itu, K_1 mengerjakan perangkat tes P dan kemudian perangkat tes Q, serta K_2 mengerjakan perangkat tes Q dan kemudian perangkat tes P. Setiap kelompok mengerjakan dua perangkat tes dalam urutan yang terbalik. Selanjutnya, dari hasil pengerjaan itu ditemukan patokan konversi.

Pada rancangan C, kelompok peserta tes tidak dibentuk atas dasar acak. Pada rancangan ini, peserta tes hanya disyaratkan ada dua kelompok, yaitu K_1 dan K_2 . Sekalipun demikian jika kelompok peserta tes diperoleh dengan cara mengacak tidak menjadi masalah. Selanjutnya perangkat tes P dan Q masing-masing ditambah dengan perangkat tes pengait atau perangkat tes jangkar atau anchor (R), sehingga perangkat tes menjadi P+R dan Q+R. Setelah itu K_1 mengerjakan perangkat tes P+R dan K_2 mengerjakan perangkat tes Q+R (lihat Tabel 1), yang akhirnya ditemukan formula konversi.

Jika ketiga rancangan tersebut di atas dibandingkan, maka ditemukan beberapa hal. Dari segi kemantapan, rancangan B adalah yang paling mantap karena setiap kelompok mengerjakan semua perangkat tes, dan sebaliknya setiap perangkat tes dikerjakan setiap kelompok peserta tes. Berikutnya rancangan C, karena setiap kelompok mengerjakan sebagian perangkat tes yang sama. Sedangkan yang terakhir adalah rancangan A, karena rancangan ini hanya bertopang pada kesetaraan peserta tes. Tetapi dalam hal mempunyai tes yang panjang dan mempunyai banyak peserta, rancangan A paling praktis untuk digunakan, karena setiap kelompok hanya mengerjakan satu perangkat tes. Dilihat dari penyiapan peserta tes, rancangan C

adalah yang paling praktis, karena rancangan ini tidak mensyaratkan kesetaraan kelompok peserta tes. Dengan demikian, dengan mempertimbangkan kepraktisan, maka dalam penelitian ini digunakan rancangan C.

Proses penyetaraan boleh dipandang sebagai suatu prosedur untuk menetapkan konversi skor yang berasal dari perangkat tes yang berbeda. Jika dua perangkat tes P dan Q telah disetarakan dan seseorang mengambil perangkat tes P, skor P dapat diekuivalensikan pada skala skor dari perangkat tes Q. Fungsi hubungan skor P untuk setara dengan skor Q dinyatakan melalui persamaan $Q = f(p)$. Skor tes seseorang yang mengambil P dapat dibandingkan dengan skor seseorang yang mengambil tes Y melalui persamaan transformasi.

Dalam pelaksanaannya, dikenal dua macam cara penyetaraan yakni cara linear dan ekuipersentil. Cara linear menghubungkan skor konversi dengan skor asalnya melalui suatu fungsi linear, sedangkan cara ekuipersentil akan mengkonversi skor berdasarkan persentil yang sama. Selanjutnya dalam penelitian ini dibatasi pada penyetaraan dengan cara linear saja.

5. Penyetaraan metode linear rancangan C.

Metode penyetaraan linear dilakukan dengan prosedur regresi linear. Konversi skor perangkat tes P ke skor perangkat tes Q adalah skor Q^* , dan sebaliknya konversi skor perangkat tes Q ke skor perangkat tes P adalah skor P^* . Pada konversi metode linear, skor Q^* merupakan fungsi linear dari skor tes P, atau sebaliknya skor P^* merupakan fungsi linear dari skor Q. Menurut Crocker & Algina (1986) hubungan linear skor perangkat tes Q^* dan skor perangkat P, dirumuskan sebagai berikut.

$$Q^* = a(X - c) + d. \dots (1)$$

Keterangan: a, c, d lihat pada uraian berikut.

Pada rancangan C ini, K_1 mengerjakan perangkat tes P dan R, sedangkan K_2 mengerjakan perangkat tes Q + R. Perangkat tes R yang dikerjakan oleh K_1

menghasilkan skor R_1 dengan rerata μ_{R1} dan simpangan baku σ_{R1} . Perangkat tes R yang dikerjakan K_2 menghasilkan skor R_2 dengan rerata μ_{R2} dan simpangan baku σ_{R2} . Hasil gabungan perangkat tes yang dikerjakan oleh kelompok K_1 dan K_2 adalah skor R dengan rerata μ_R dan simpangan baku σ_R .

Perangkat tes P hanya dikerjakan oleh K_1 dan menghasilkan skor P dengan rerata μ_P dan simpangan baku σ_P , sedangkan perangkat tes Q hanya dikerjakan oleh K_2 dan menghasilkan skor Q dengan rerata μ_Q dan simpangan baku σ_Q . Regresi diperoleh dari skor kedua perangkat tes (P dan Q) terhadap skor R. Dari kedua regresi tersebut diperoleh koefisien arah b_{PR1} dan b_{QR2} .

Dengan demikian formula statistik penyetaraan linear rancangan C dari perangkat tes P dan Q dinyatakan sebagai berikut.

$$\frac{P - \{ \mu_P + b_{PR1} (\mu_R - \mu_{R1}) \}}{\sqrt{\sigma_P^2 + b_{PR1}^2 (\sigma_R^2 - \sigma_{R1}^2)}} = \frac{Q - \{ \mu_Q + b_{QR2} (\mu_R - \mu_{R2}) \}}{\sqrt{\sigma_Q^2 + b_{QR2}^2 (\sigma_R^2 - \sigma_{R2}^2)}} \quad \dots(2)$$

Berdasarkan rumus (1) dan (2) didapatkan formula konversi:

$$Q^* = a (P - c) + d,$$

Dengan nilai a, c, dan d sebagai berikut:

$$a = \sqrt{\frac{\sigma_Q^2 + b_{QR2}^2 (\sigma_R^2 - \sigma_{R2}^2)}{\sigma_P^2 + b_{PR1}^2 (\sigma_R^2 - \sigma_{R1}^2)}}$$

$$c = \mu_P + b_{PR1} (\mu_R - \mu_{R1})$$

$$d = \mu_Q + b_{QR2} (\mu_R - \mu_{R2}) \quad (\text{diadaptasi dari Crocker \& Algina, 1986}).$$

Sehingga rumus penyetaraan dari skor perangkat tes P ke skor perangkat Q untuk rancangan C dapat pula dinyatakan rumus (3) berikut.

$$Q^* = \left\{ \sqrt{\frac{\sigma^2_Q + b^2_{QR2} (\sigma^2_R - \sigma^2_{R2})}{\sigma^2_P + b^2_{PR1} (\sigma^2_R - \sigma^2_{R1})}} \right\} \left\{ P - (\mu_P + b_{PR1} (\mu_R - \mu_{R1})) \right\} + \left\{ (\mu_Q + b_{QR2} (\mu_R - \mu_{R2})) \right\}$$

Dengan cara yang sama, dapat pula dibuat rumus penyetaraan dari skor perangkat tes Q ke skor perangkat P (kebalikannya), seperti rumus (4) berikut.

$$P^* = \left\{ \sqrt{\frac{\sigma^2_P + b^2_{PR1} (\sigma^2_R - \sigma^2_{R1})}{\sigma^2_Q + b^2_{QR2} (\sigma^2_R - \sigma^2_{R2})}} \right\} \left\{ Q - (\mu_Q + b_{QR2} (\mu_R - \mu_{R2})) \right\} + \left\{ (\mu_P + b_{PR1} (\mu_R - \mu_{R1})) \right\}$$

Nilai statistik dari data skor yang diperlukan pada penyetaraan linear rancangan C, seperti pada Tabel 2 berikut.

Tabel 2. Nilai Statistik yang Dibutuhkan pada Penyetaraan Perangkat Tes Rancangan C

Kelompok	Nilai Statistik	Perangkat Tes		
		P	Q	R
K ₁	μ	YA	-	YA
	σ	YA	-	YA
	b _{PR1}	YA	-	-
K ₂	μ	-	YA	YA
	σ	-	YA	YA
	b _{QR2}	-	-	-
K ₁ + K ₂	μ	-	-	YA
	σ	-	-	YA

Keterangan: Diadaptasi dari Crocker & Algina, 1986.

6. Standar kesalahan penyetaraan

Pada pengukuran hasil belajar, salah satu kriteria baik tidaknya alat ukur yang digunakan didasarkan kepada standar kesalahan pengukuran (SEM). Semakin kecil nilai kesalahan pengukuran suatu tes, semakin dapat diandalkan atau reliabel tes

tersebut. Apabila kesalahan pengukuran yang terjadi sama dengan nol, berarti skor yang diperoleh peserta tes adalah skor yang sebenarnya, dan ini menurut Allen & Yen (1979) alat ukur tersebut telah mencapai suatu kondisi yang ideal.

Demikian pula halnya pada penyetaraan skor beberapa perangkat tes, juga terdapat standar kesalahan penyetaraan atau *Standard Error of Equating* (SEE). Penyetaraan yang baik adalah penyetaraan yang memiliki standar kesalahan kecil. Standar kesalahan penyetaraan (SEE) didasarkan pada variansi dan rerata skor dari kedua perangkat tes yang disetarakan. Kesalahan pengukuran dalam penyetaraan disebabkan oleh kesalahan dalam pengambilan sampel peserta tes. Ketidaktepatan dalam pengambilan sampel peserta tes dapat menyebabkan terjadinya bias dalam rerata dan simpangan baku skor kedua perangkat tes yang disetarakan.

Petersen, Kolen, dan Hoover (dalam Linn, 1989) mengemukakan bahwa nilai SEE dapat ditentukan dari perangkat tes P ke perangkat tes Q, atau sebaliknya dari perangkat tes Q ke perangkat tes P. SEE dari perangkat tes P ke perangkat tes Q dinyatakan sebagai berikut.

$$SEE[Q(P_0)] = \{ [(\sigma_Q^2 / 2)] [2 + (P_0 - \mu_P)^2 / \sigma_P^2] [(1/N_P) + (1/N_Q)] \} \dots (5)$$

Keterangan:

$SEE[Q(P_0)]$ = Standar kesalahan penyetaraan (P ke Q)

σ_P^2 = Varians skor perangkat tes P

σ_Q^2 = Varians skor perangkat tes Q

P_0 = Skor perangkat tes P

μ_P = Rerata skor perangkat tes P

N_P = Jumlah peserta tes yang merespons perangkat tes P

N_Q = Jumlah peserta tes yang merespons perangkat tes Q

Pada dasarnya penentuan kesalahan pengukuran penyetaraan disyaratkan agar skor dari perangkat tes yang disetarakan harus berdistribusi normal. Namun, mengingat asumsi normalitas sulit ditemui dalam pelaksanaan, untuk sampel cukup besar persyaratan normalitas tidak dipersyaratkan secara ketat (Salasi, 1997).

Penyetaraan dengan sampel besar dapat mengurangi atau menekan kesalahan pengukuran penyetaraan menjadi kecil, dan dapat diabaikan.

BAB III

METODOLOGI PENELITIAN

Matematika sebagai suatu bidang ilmu, diajarkan di hampir setiap jenis dan jenjang pendidikan, sejak Sekolah Dasar hingga Perguruan Tinggi. Peran dan pentingnya matematika dalam berbagai ilmu dan juga dalam kehidupan sehari-hari cukup besar dan tidak dapat diabaikan. Di lain pihak prestasi mahasiswa cenderung kurang baik, karena pada umumnya matakuliah yang banyak mengandung unsur matematika termasuk dalam matakuliah yang sukar, sehingga dalam pelaksanaannya bagi cukup banyak mahasiswa mengambil jalan pintas dengan mencari bocoran soal atau kunci. Oleh karena itu, dalam penelitian ini perangkat tes yang dikembangkan untuk kebutuhan penyetaraan tes adalah matakuliah Matematika dengan kode matakuliah PGSD2204 yang terdapat pada Program Penyetaraan D2 PGSD.

A. Tempat dan Waktu Penelitian

Penelitian ini dilakukan terhadap mahasiswa Program D2 PGSD yang berada di UPBJJ Bogor dan UPBJJ Lampung. Pelaksanaan penelitian berlangsung selama kurang lebih enam bulan, sejak Mei sampai dengan akhir November 2000.

B. Instrumen Penelitian

Untuk kepentingan penyetaraan dalam penelitian ini diperlukan dua perangkat tes sebagai instrumen penelitian, yaitu perangkat tes tipe I dan tipe II. Perangkat tes tipe I dibentuk menjadi perangkat tes P ditambah dengan perangkat tes anchor R. Perangkat tes tipe II dibentuk menjadi perangkat tes Q ditambah perangkat tes anchor R. Kedua tipe perangkat tes yang digunakan dikembangkan dari kisi-kisi yang sama, yaitu kisi-kisi yang dikembangkan oleh UPS FKIP. Kemudian setiap butir dikembangkan secara berpasangan.

Kedua perangkat tes (tipe I dan II) masing-masing terdiri dari 49 butir soal bentuk pilihan ganda (objektif) dengan empat alternatif jawaban. Dari 49 butir soal tersebut, 39 butir bentuk objektif biasa (tipe A) dan 10 butir bentuk objektif kompleks (tipe D). Kemudian untuk kepentingan penyetaraan dengan rancangan C ini, banyaknya butir pada perangkat tes P dan Q masing-masing 43 butir soal, dan perangkat tes anchor (R) sebanyak 6 butir soal.

C. Populasi dan Sampel Penelitian

Seperti telah disebutkan pada pembatasan penelitian, bahwa penelitian ini diadakan terhadap mahasiswa yang merespons matakuliah Matematika (PGSD 2204). Dengan demikian, sebagai populasi dalam penelitian ini adalah mahasiswa Program D2 PGSD yang mengambil matakuliah Matematika (PGSD 2204). Pengambilan matakuliah ini dengan alasan: (1) matakuliah ini lebih banyak direspons mahasiswa, (2) matakuliah ini termasuk kategori sukar (tingkat kelulusannya sangat rendah), dan (3) karena matakuliah ini termasuk kategori matakuliah yang sukar, maka cenderung terjadi kebocoran atau penyimpangan lainnya dalam penyelenggaraan ujian.

Kemudian, untuk kepentingan penelitian ini sampel mahasiswa diambil dari dua UPBJJ, yaitu UPBJJ Bogor dan UPBJJ Bandar Lampung. Selanjutnya, untuk UPBJJ Bogor diambil Pokjar yang berada di Kabupaten/Kotamadya Tangerang, dan untuk UPBJJ Bandar Lampung diambil Pokjar yang berada di Kabupaten Lampung Utara (Kotabumi). Masing-masing Pokjar jumlah mahasiswa yang terjaring sebagai sampel secara berurutan diperoleh 68 dan 71 mahasiswa. Dengan demikian, teknik pengambilan sampel yang digunakan dalam penelitian ini adalah *Purposive Cluster Sampling*.

D. Metode Pengumpulan Data

Sesuai dengan rumusan masalah dan tujuan penelitian, data yang dikehendaki dalam penyetaraan perangkat tes adalah skor hasil pengerjaan mahasiswa terhadap tes matakuliah Matematika (PGSD 2204). Data tersebut dikumpulkan atau diperoleh

melalui tes terhadap dua kelompok mahasiswa (kelompok K_1 dan K_2). Mahasiswa kelompok K_1 , yaitu kelompok mahasiswa dari UPBJJ Bandar Lampung (Pokjar Kotabumi) diberi perangkat tes tipe I (perangkat $P + R$), dan mahasiswa kelompok K_2 , yaitu kelompok mahasiswa dari UPBJJ Bogor (Pokjar Tangerang) diberi perangkat tes tipe II (perangkat $Q + R$).

E. Metode Analisis Data

Analisis data dilakukan dengan menempuh dua tahapan, yaitu pertama analisis butir untuk menentukan karakteristik tes, dan kedua analisis statistik deskriptif untuk kepentingan penyetaraan tes. Pada analisis butir, data yang dibutuhkan berupa respons mahasiswa terhadap tes yang diberikan. Hasil analisis butir terhadap kedua perangkat tes tersebut berupa: reliabilitas, *SEM*, indeks kesukaran, indeks daya beda tes, dan skor mahasiswa dalam skala 49 (banyak butir). Analisis butir ini menggunakan program *MicroCAT*. Selanjutnya, skor mahasiswa diolah kembali menggunakan statistik deskriptif untuk menentukan nilai statistik μ , σ^2 , σ , serta nilai lainnya yang dibutuhkan (seperti pada Tabel 2) dari masing-masing perangkat tes.

BAB IV

HASIL PENELITIAN DAN PEMBAHASAN

A. Hasil Penelitian

Sebelum melaksanakan tes untuk kepentingan penyetaraan terlebih dahulu dilakukan penelaahan terhadap tes yang digunakan. Setelah perangkat tes digunakan didapat data respons peserta tes terhadap kedua perangkat tes yang telah dirancang (tes tipe I dan tes tipe II). Tes tipe I mengandung perangkat tes P dan R, dan tes tipe II mengandung perangkat tes Q dan R (R merupakan perangkat tes pengait atau anchor). Kemudian, data berupa respons peserta tes dianalisis menggunakan program *MicroCAT*, diperoleh data tentang karakteristik tes dan data untuk kepentingan penyetaraan tes. Dengan demikian, pembahasan hasil penelitian ini meliputi analisis tentang karakteristik tes dan penyetaraan perangkat tes, sesuai dengan tujuan penelitian yang telah dirumuskan.

1. Karakteristik Tes

Seperti telah dikemukakan bahwa perangkat tes yang digunakan terdiri dari tiga perangkat tes, yaitu perangkat tes P, Q, dan perangkat tes R, yang masing-masing perangkat tes terdiri 40 butir soal untuk perangkat tes P dan Q, dan enam butir soal untuk perangkat tes R. Hasil analisis butir tes secara lengkap dapat dilihat pada Tabel 3 berikut.

Tabel 3. Karakteristik Tes Perangkat Tes P,Q, dan R.

Perangkat Tes	Karakteristik Tes			
	Indeks Kesukaran	Indeks Daya Beda	Reliabilitas	SEM
P	0,388	0,416	0,785	2,863
Q	0,522	0,375	0,746	2,788
R	0,333	0,584	0,162	0,996

Keterangan: Skala skor adalah 0 - 43 untuk perangkat tes P dan Q, skala 0 – 6 untuk perangkat tes R.

Berdasarkan Tabel 3 di atas, Indeks kesukaran ketiga perangkat tes (P, Q, dan R) masing-masing secara berurut 0,388; 0,522; dan 0,333. Indeks daya beda masing-masing perangkat tes sebesar 0,416; 0,375; dan 0,584. Indeks kesukaran dan indeks daya beda tes diperoleh dengan cara merata-ratakan seluruh indeks kesukaran dan indeks daya beda butir yang membentuk masing-masing perangkat tes tersebut. Angka-angka indeks kesukaran soal di atas menunjukkan bahwa ketiga perangkat tes tersebut memiliki tingkat kesukaran sedang, dan tingkat kesukaran perangkat tes P lebih sukar dibandingkan tingkat kesukaran perangkat tes Q. Daya beda ketiga perangkat tes secara keseluruhan melebihi 0,30 (tergolong baik).

Selain tingkat kesukaran dan daya beda, perlu diperhatikan besarnya koefisien reliabilitas dan besarnya kesalahan pengukuran dari ketiga perangkat tes yang digunakan. Secara berurutan, perangkat tes P, Q, dan R, masing-masing koefisien reliabilitasnya adalah 0,785; 0,746; 0,162. Sedangkan nilai kesalahan pengukuran masing-masing perangkat tes sebesar 2,863; 2,788; 0,996. Untuk perangkat tes P dan Q, skala kesalahan pengukurannya antara 0-43, sementara untuk perangkat tes R skala pengukurannya antara 0-6.

2. Penyetaraan Perangkat Tes

Seperti telah dikemukakan di atas, perangkat tes yang digunakan dalam penelitian ini mengikuti rancangan penyetaraan perangkat tes bentuk C, sehingga perangkat tesnya ada dua tipe, yaitu tipe I dan tipe II. Tes tipe I dibentuk oleh dua perangkat tes, yakni perangkat tes P dan R. Tes tipe II dibentuk oleh perangkat tes Q dan R. Dengan demikian sebenarnya perangkat tes yang digunakan ada tiga perangkat tes, yaitu perangkat tes P, Q, dan R. Dengan kata lain, perangkat tes R digabungkan di dua perangkat tes lainnya, atau perangkat tes R merupakan tes anchor yang sama pada kedua tipe, sedangkan P dan Q merupakan perangkat tes yang akan disetarakan. Jumlah butir masing-masing perangkat tes sebesar 43 butir untuk perangkat tes P dan

Q, dan enam butir untuk perangkat tes R. Dengan demikian, jumlah butir masing-masing tipe perangkat tes sebesar 49 butir.

Hasil analisis terhadap hasil respons peserta tes, diperoleh data skor [eserta tes yang diperlukan untuk kepentingan penyetaraan. Skor peserta tes pada ketiga perangkat tes dinyatakan dalam skala 0 – 43, dengan pertimbangan untuk keseragaman dan memudahkan dalam penafsiran skor dari ang-angka statistik yang dihasilkan.

Rumus konversi skor dari P ke Q (perangkat tes P sebagai pembanding) seperti telah dikemukakan pada Bab II, yaitu $Q^* = a (P - c) + d$, dengan: a, c, dan d sebagai berikut:

$$a = \sqrt{\frac{\sigma^2_Q + b^2_{QR2}(\mu^2_R - \mu^2_{R2})}{\sigma^2_P + b^2_{PRI}(\mu^2_R - \mu^2_{R1})}}$$

$$c = \mu_P + b_{PRI}(\mu_R - \mu_{R1})$$

$$d = \mu_Q + b_{QR2}(\mu_R - \mu_{R2}) \qquad \text{(diadaptasi dari Crocker \& Algina, 1986).}$$

Data nilai statistik hasil analisis skor peserta tes pada perangkat tes P, Q, dan R dirangkum pada Tabel 4 berikut.

Tabel 4. Data Nilai Statistik Skor Peserta Tes Pada Perangkat Tes P,Q, dan R.

Kelompok	Nilai Statistik	Perangkat Tes		
		P	Q	R
K ₁	μ	16,691	-	11,488
	σ	6,172	-	7,880
	b _{PRI}	0,588	-	-
K ₂	μ	-	22,451	14,032
	σ	-	5,528	7,740
	b _{QR2}	-	0,344	-
K ₁ + K ₂	μ	-	-	12,893
	σ	-	-	7,792

Keterangan: Skor pada setiap perangkat tes dalam 0-43, dengan pembulatan tiga desimal.

Berdasarkan nilai statistik pada Tabel 4 diperoleh nilai a , b , dan d , masing-masing sebesar $a = 0,909$; $c = 17,517$; dan $d = 22,059$. Angka-angka tersebut dimasukkan ke dalam persamaan $Q^* = a(P - c) + d$, sehingga diperoleh persamaan konversi tes matakuliah Matematika (PGSD 2204) dari perangkat tes P ke perangkat tes Q , yaitu $Q^* = 0,909(P - 17,517) + 22,059$ atau $Q^* = 0,909P + 6,136$.

Memerperhatikan persamaan regresi yang diperoleh, skor pada kedua perangkat tes terdapat perbedaan kira-kira enam angka pada skala 43. Artinya, skor yang diperoleh seseorang peserta tes dengan mengikuti perangkat tes P dan skor peserta tes yang mengikuti perangkat tes Q terdapat perbedaan. Misalnya, skor pada perangkat tes $P = 11$, sama kedudukannya dengan skor pada perangkat $Q = 16,135$ (dalam skala 0-43); skor pada perangkat tes $P = 25$, sama kedudukannya dengan skor pada perangkat $Q = 28,861$; skor pada $P = 35$, sama kedudukannya dengan skor $Q = 37,951$, dan seterusnya.

Sebaliknya, dengan cara yang sama dapat ditentukan pula persamaan konversi skor dari perangkat tes Q ke perangkat tes P (Q sebagai pembanding, yaitu $P^* = 1,100(Q - 22,059) + 17,517$, atau $P^* = 1,100Q - 6,748$. Persamaan ini sama dengan persamaan yang diperoleh dari invers persamaan konversi skor penyetaraan perangkat tes P ke Q ($Q^* = 0,909P + 6,136$). Misal, skor pada $Q = 16,135$, akan sama dengan skor pada $P = 11$. Perbedaan nilai pada kedua model jika saling dikonversikan pada kedua persamaan skor penyetaraan hanya berbeda dalam pembulatan dua atau tiga angka desimal.

3. Kesalahan Pengukuran Penyetaraan

Seperti dikemukakan pada Bab II, untuk menentukan kesalahan pengukuran penyetaraan digunakan rumus:

$$SEE[Q(P_0)] = \{ [(\sigma_Q^2 / 2)] [2 + (P_0 - \mu_P)^2 / \sigma_P^2] [(1/N_P) + (1/N_Q)] \}^{1/2}.$$

Untuk kepentingan perhitungan kesalahan pengukuran penyetaraan ini diperlukan nilai statistik seperti pada Tabel 5 berikut.

Tabel 5. Nilai Statistik untuk Menentukan Kesalahan Pengukuran Penyetaraan Perangkat Tes P dan Q

Nilai Statistik	Perangkat Tes	
	P	Q
μ	16,691	22,451
σ^2	38,094	30,559
N	68	71

Nilai statistik pada Tabel 5 di atas, disubstitusikan ke dalam rumus kesalahan penyetaraan, sehingga diperoleh besarnya kesalahan pengukuran penyetaraan (SEE) dari perangkat tes P ke perangkat tes Q, dan juga sebaliknya.

Besarnya kesalahan pengukuran penyetaraan dari perangkat tes P ke perangkat tes Q untuk tes matakuliah Matematika (PGSD 2204), berdasarkan rumus *SEE* dan nilai statistik yang tertera pada Tabel 5 diperoleh:

$SEE[Q(P_0)] = \{ [0,440] [2 + (P_0 - 16,691)^2 / 38,094] \}^{1/2}$, sehingga diperoleh nilai-nilai $SEE[Q(P_0)]$ setiap skor serta nilai rata-rata kesalahan pengukuran penyetaraan seperti pada Tabel 6.

Dengan cara yang sama, kesalahan pengukuran penyetaraan dari perangkat Q ke perangkat tes P untuk tes matakuliah Matematika (PPDG 2204) diperoleh:

$SEE[P(Q_0)] = \{ [0,564] [2 + (Q_0 - 22,451)^2 / 30,559] \}^{1/2}$, sehingga untuk setiap nilai Q_0 tertentu akan diperoleh nilai $SEE[P(Q_0)]$ dan nilai reratanya, seperti terlihat pada Tabel 7.

Tabel 6. Kesalahan Pengukuran Penyetaraan dari Perangkat Tes P ke Perangkat Tes Q.

Skor P	F(P)	$SEE[Q(P_0)]$	$F(P).SEE[Q(P_0)]$
8	1	1,324	1,324
9	3	1,250	3,750
10	8	2,182	9,456
11	4	1,120	4,479
12	6	1,065	6,390
13	4	1,019	4,074
14	6	0,964	5,782
15	3	0,955	2,867
16	3	0,866	2,657
17	3	0,939	2,816
18	5	0,949	4,743
19	1	0,970	0,970
20	3	1,003	4,013
21	1	1,046	1,046
22	3	1,098	3,294
23	2	1,157	2,315
25	5	1,295	6,476
26	2	1,371	2,743
30	1	1,712	1,711
31	1	1,801	3,603
32	2	1,894	3,788
Σ	68		78,293
Rata-rata $SEE[Q(P_0)]$			1,151

Keterangan:

$F(P)$ = Frekuensi skor pada perangkat P, \forall P tertentu.

$SEE[Q(P_0)]$ = Standar kesalahan penyetaraan dari perangkat P ke Q, pada $P=P_0$.

Tabel 7. Kesalahan Pengukuran Penyetaraan dari Perangkat Tes Q ke Perangkat Tes P.

Skor Q	F(Q)	$SEE[P(Q_0)]$	$F(P).SEE[P(Q_0)]$
10	1	1,968	1,968
11	1	1,857	1,857
12	1	1,748	1,748
13	2	1,642	3,285
14	2	1,542	3,083
15	1	1,446	1,446
16	5	1,357	6,787
17	3	1,276	3,829
18	1	1,205	1,205
19	7	1,144	8,011
20	1	1,097	1,097
21	5	1,064	5,324
22	4	1,049	4,195
23	2	1,049	2,099
24	4	1,067	4,269
25	8	1,101	8,809
26	3	1,150	3,449
27	4	1,211	4,845
28	7	1,284	8,987
29	5	1,366	6,828
31	1	1,551	1,551
32	2	1,653	3,305
33	1	1,758	1,758
Σ	71		89,735
Rata-rata $SEE[Q(P_0)]$			1,263

Keterangan:

$F(Q)$ = Frekuensi skor pada perangkat Q, \forall Q tertentu.

$SEE[P(Q_0)]$ = Standar kesalahan penyetaraan dari perangkat Q ke P, pada $Q=Q_0$.

Hasil perhitungan seperti tertera pada Tabel 6 dan Tabel 7 menunjukkan, besarnya nilai rerata kesalahan pengukuran penyetaraan pada kedua persamaan konversi hasil penyetaraan perangkat tes matakuliah Matematika (PGSD 2204), yaitu masing-masing sebesar $SEE[Q(P_0)] = 1,151$ dan $SEE[Q(P_0)] = 1,263$. Jika dibandingkan terhadap skala yang digunakan (0-43), perbedaan angka kedua nilai

rerata kesalahan pengukuran penyetaraan (SEE) tidak terlalu besar, yakni 0,112 (2,8%), atau perbedaannya tidak cukup berarti.

B. Pembahasan Hasil Penelitian

Berdasarkan temuan penelitian dari hasil analisis data yang telah disebutkan di atas, dapat dikemukakan penjelasan dan pembahasan seperti berikut.

1. Karakteristik Tes

Kualitas perangkat tes yang digunakan sebagai perangkat tes yang disetarakan secara keseluruhan dapat dikatakan perangkat tes yang baik untuk digunakan sebagai alat ukur matakuliah Matematika (PGSD 2204). Rerata indeks kesukaran (indeks kesukaran tes) dan rerata indeks daya beda (indeks daya beda tes) memenuhi syarat atau cukup memadai sebagai alat ukur. Indeks kesukaran tes P, Q, dan R masing-masing sebesar 0,388; 0,522; dan 0,333. Menurut Allen & Yen (1979), indeks kesukaran yang terletak antara 0,30 sampai dengan 0,70 termasuk kategori tes dengan tingkat kesukaran sedang. Dengan demikian, ketiga perangkat tes yang digunakan dalam penyetaraan tes matakuliah PGSD 2204 termasuk dalam kategori tingkat kesukaran sedang.

Indeks daya beda ketiga perangkat tes P, Q, dan R secara berurut sebesar 0,416; 0,375; dan 0,584. Dengan indeks daya beda seperti yang tertera di atas menunjukkan bahwa ketiga perangkat tes yang digunakan memenuhi persyaratan atau cukup memadai untuk dijadikan alat ukur. Artinya, perangkat tes yang digunakan memiliki kemampuan untuk membedakan antara kelompok peserta tes yang pandai (berprestasi tinggi) dengan peserta tes yang kurang pandai (prestasi rendah). Menurut Suyabrata (1999) dan Fernandes (1984) perangkat tes yang mempunyai indeks daya beda lebih besar dari 0,30 sudah cukup baik untuk digunakan sebagai alat ukur.

Koefisien reliabilitas ketiga perangkat tes P, Q, dan R secara berurut sebesar 0,785; 0,746; dan 0,162. Dilihat dari besarnya koefisien reliabilitasnya, dari ketiga

perangkat tes tersebut yang memenuhi syarat hanya perangkat tes P dan Q. Hal ini sesuai dengan pendapat Gronlund (1981), bahwa koefisien reliabilitas tes buatan guru dapat dikaitkan baik bila besarnya koefisien reliabilitas sebesar 0,60 sampai dengan 0,85. Sedangkan koefisien reliabilitas untuk perangkat tes R hanya sebesar 0,162. Secara sepintas koefisien reliabilitas tes R sangat rendah bila dibandingkan dengan koefisien reliabilitas tes P dan Q. Tetapi, sebenarnya tidak demikian, karena jumlah butir perangkat tes R hanya enam butir soal, sementara perangkat tes P dan Q masing-masing 43 butir soal. Menurut Azwar (1999) koefisien reliabilitas sangat dipengaruhi oleh panjang tes (jumlah butir soal). Bila banyaknya butir diperbanyak 4 kali lipat, maka koefisien reliabilitas akan naik menjadi 2,5 kali koefisien reliabilitas semula. Dengan demikian, bila banyak butir pada perangkat tes R disejajarkan dengan perangkat tes P dan Q (tujuh kali banyak butir tes semula), maka prediksi koefisien reliabilitas tes R sebesar 0,709. Artinya, perangkat tes R dapat dipercaya untuk digunakan sebagai alat ukur.

Selanjutnya, dilihat dari nilai standar kesalahan pengukuran, secara berurutan perangkat tes P, Q, dan R, masing-masing sebesar 2,863; 2,788; dan 0,996. Berdasarkan angka-angka tersebut, seluruh perangkat tes memenuhi kriteria standar kesalahan pengukuran minimal, yaitu seperti yang diungkapkan oleh Hopkin *et al.*, yaitu perkiraan akurasi dari standar kesalahan pengukuran seharusnya di bawah harga $\sigma_e = 0,43\sqrt{k}$, dimana k = banyaknya butir soal. Untuk $k = 43$, maka harga $\sigma_e = 2,820$; dan untuk $k = 6$, maka harga $\sigma_e = 1,053$. Dengan demikian kesalahan pengukuran dari ketiga perangkat tes yang digunakan masih dalam kesalahan pengukuran yang wajar. Artinya, perangkat tes yang digunakan dapat dipercaya sebagai alat ukur, dan hasil ukurnya atau skor peserta tes mendekati skor yang sebenarnya.

2. Penyetaraan Perangkat Tes

Dalam penyetaraan perangkat tes telah ditemukan persamaan konversi skor dari perangkat tes P ke perangkat tes Q, yaitu $Q^* = 0,909P + 6,136$, dan persamaan konversi dari perangkat Q ke perangkat P, yaitu $P^* = 1,100Q - 6,748$. Secara empiris telah ditunjukkan pasangan skor antara perangkat tes P dan perangkat tes Q pada kedua persamaan konversi. Persamaan konversi yang diperoleh menggambarkan bahwa skor pada P dan Q memiliki perbedaan yang cukup berarti. Misalnya untuk $P = 12,000$; maka skor pada $Q = 17,044$. demikian juga sebaliknya, untuk $Q = 21,000$; maka skor pada $P = 16,352$. Artinya, peserta tes yang menggunakan perangkat tes P (peserta tes di Lampung), skor yang diperolehnya sebesar 12,000 dikonversikan kepada perangkat tes Q (yang digunakan di Tangerang) setara dengan skor 17,044. Sebaliknya, peserta tes yang menggunakan perangkat tes Q (peserta tes di Tabgerang), skor yang diperolehnya sebesar 21,000 dikonversikan kepada perangkat tes P (yang digunakan di Lampung) setara dengan skor 16,352.

Perbedaan skor yang cukup berarti dari hasil konversi Q^* maupun P^* , juga dapat diartikan bahwa kedua perangkat tes memiliki tingkat kesukaran yang berbeda pula. Bila dilihat dari indeks kesukaran tes antara perangkat tes P dan Q, perangkat tes Q lebih mudah dibandingkan dengan perangkat tes P (indeks kesukaran perangkat tes P = 0,388 dan Q = 0,522), sehingga sangat wajar rerata skor peserta tes yang mengambil perangkat tes P (peserta tes di Lampung) lebih rendah di bandingkan rerata skor peserta tes yang mengambil perangkat tes Q (peserta tes di Tangerang).

Kesalahan pengukuran dalam proses penyetaraan dua perangkat tes perlu mendapat pertimbangan untuk menentukan bentuk persamaan konversi mana yang terbaik, dengan melihat rerata kesalahan yang paling kecil. Hasil perhitungan seperti tertera pada Tabel 6 dan Tabel 7, pada kedua persamaan konversi penyetaraan perangkat tes memiliki nilai kesalahan pengukuran penyetaraan $Q^* = 0,909P + 6,136$ adalah $SEE[Q(P_0)] = 1,151$ dan $P^* = 1,100Q - 6,748$ adalah $SEE[Q(P_0)] = 1,263$. Perbedaan rerata SEE kedua persamaan konversi sangat kecil, yaitu 0,112 pada skala

0-43. Namun demikian, disarankan persamaan konversi yang digunakan adalah $Q^* = 0,909P + 6,136$, karena memiliki rerata *SEE* terkecil, yaitu 1,151.

BAB V

KESIMPULAN DAN SARAN

A. Kesimpulan

Berdasarkan temuan hasil penelitian dan pembahasan, dapat dikemukakan kesimpulan sebagai berikut.

Persamaan atau formula konversi skor dari dua perangkat tes matakuliah Matematika (PGSD 2204), yaitu antara perangkat tes P dan perangkat tes Q adalah $Q^* = 0,909P + 6,136$ atau $P^* = 1,100Q - 6,748$. Kesalahan pengukuran penyetaraan pada masing-masing persamaan konversi sebesar $SEE[Q(P_0)] = 1,151$ dan $SEE[P(Q_0)] = 1,263$. Berdasarkan kedua persamaan konversi tersebut tergambar bahwa kedua perangkat tes memiliki tingkat kesukaran yang berbeda. Perangkat tes P lebih sukar dibandingkan dengan perangkat tes Q. Dengan demikian, persamaan konversi hasil penyetaraan yang terbaik dengan rerata kesalahan pengukuran penyetaraan terkecil adalah $Q^* = 0,909P + 6,136$.

B. Saran

Mengingat pentingnya penyetaraan perangkat tes dalam penilaian hasil belajar, saat ini, khususnya UT perlu mengembangkan perangkat tes yang setara untuk setiap matakuliah. Oleh karena itu, kepada pihak-pihak yang terlibat dalam pengembangan tes, dianjurkan untuk mengupayakan terciptanya beberapa perangkat tes yang telah disetarakan. Beberapa perangkat tes yang memiliki persamaan konversi tertentu dapat digunakan untuk menilai hasil belajar peserta tes /mahasiswa secara adil. Di samping itu, dengan adanya perangkat tes yang setara dapat digunakan untuk mengatasi bila terjadi kebocoran soal dalam penyelenggaraan ujian, dan juga berguna untuk memantau kemajuan prestasi pendidikan di berbagai daerah atau dari waktu ke waktu.

DAFATAR PUSTAKA

- Allen, M.J. and Yen, W.M. (1979). *Introduction to measurement theory*. California: Brooks/Cole Publishing Company.
- Linn, R.L. (1989) (Ed.) *Educational measurement*. New York: Macmillan Publishing Company.
- Crocker, L. And Algina, J. (1986). *Introduction to classical and modern test theory*. New York: Holt, Rinehart and Winston, Inc.
- Depdiknas. (2000). *Sambutan Rektor pada Dies XVI Universitas Terbuka, 4 September 2000*. Jakarta: Universitas Terbuka.
- Fernandes, H.J.X. (1984). *Testing and measurement*. Jakarta: NEP, Evaluation and Curriculum Devolepment.
- Gronlund, E.N. (1982). *Constructing achievement tests*. E.C.: Prentice Hall, Inc.
- Hopkins, K.D., Stanley, J.C., and Hopkins, B.R. (1990). *Educational and psychological measurement and evaluation*. N.J.: Prentice Hall, Inc.
- Hamblethton, R.K. and Swaminathan, H. (1985). *Item response theory: Principles and applications*. Boston, M.A.: Kluwer Academic Publisher.
- Hamblethton, R.K. and Swaminathan, H. , and Roger, H.J. (1991) *Fundamentals of item response theory*. London: Sage Publications.
- Holland, P.W. and Rubin, D.B. (1982). *Test equating*. New York: Academic Press.
- Mardapi, D. (1991). *Konsep dasar teori respons butir: Perkembangan dalam bidang pengukuran pendidikan*. Yogyakarta: Cakrawala Pendidikan, nomor 3 tahun X.
- Naga, D.S. (1992). *Pengantar teori sekor pada pengukuran pendidikan*. Jakarta: Gunadarma.
- Salasi, R. (1997). *Penyetaraan perangkat tes matematika sekolah menengah pertama di Kotamadya Banda Aceh*. Yogyakarta: PPS IKIP Yogyakarta.

- Snulling, M.A. (1993). *Ekuating perangkat soal evaluasi belajar tahap akhir nasional Sekolah Dasar di Propinsi Sumatera Selatan tahun 1991/1992*. Yogyakarta: FP UGM.
- Surapranata, S. (1993). *Studi analitik tentang karakteristik internal dan formulasi ekuating perangkat tes soal EBTANAS SMP tahun pelajaran 1992/1993 di DKI Jakarta*. Yogyakarta: PPS UGM.
- Suryabrata, S. (1987). *Pengembangan tes hasil belajar*. Jakarta: CV. Raja Wali.
- _____. (1999). *Pengembangan tes hasil belajar*, Jakarta: Rajawali Press.
- Thondike, R.L. and Hagen, E.P. (1977). *Measurement and evaluation in psychology and education (ed. 4th)*. New York: John Wiley & Sons.
- Umar, J. (1991). *Memuju pembaharuan sistem pengujian pendidikan*. Jakarta: Pussisjian Depdikbud.

LAMPIRAN 1: DATA RESPONS PESERTA TES KELOMPOK I

```

KEY   DCGBCCCDCAADCACGABACBADDBBGDBBCBGBCBBCDBCBACBADBDBABDDBBABCG
OPT   44 4 4 44 44444 44 4444444 44444 4444 4444444444 444444444444
PRO   YYNNYNNYNNYYYYNNYNNYYYYNNYNNYYYYNNYNNYYYYNNYNNYYYYNNYNN
1 | CBCCAACCCDADAAABABCCBADDBBADDBCAABCBBCDBABACBADDBBADDBDDDDCA
2 | ACBCACCCDADDBADACCADCBCBCBBBDBBABAACAAAACAACCAADAACDADDCAADD
3 | ABDBACCDAAAC CCBDCACDDDBABCDACCBACCCCBABDBBDCDDBAACDBDBACCBC
4 | ABDBCCCDAAACACCCBCDCDADDBCADACCBACCCBCBDBBDCDDCAABCDDBACCBC
5 | ABDBCCDAAACCCCBDBDCBADABBCDACCBACCCACCBDBBDCDDCABACDCBBACCBC
6 | ABCBCCDAAACACCCDCDCDADDBCADACCBACCCCBDBBDCADCBACDCDBACCBC
7 | ADCAB BCDCCDAACDCDBBBBBDACDCCDADAAABBBABBBDBBCCBABBABBBAAACA
8 | CBABDADDCACBCBDBDCCDCBDCBCCDABDCCADBBACDBBCABBCDABDBDCBBDCCA
9 | CCABDBDCBABBDDCADCCDCDABBCDBBCCCBACBCCABBBBBBBADBBDBDCBBDCCA
10 | CCABDBDCBABBDDCACCCBDCDABBCDBBCCCAABBCABBBDBBBADBBDBDCBBDCCA
11 | CCBBCBDAADDCACAABACDADBBCCDDBCBABBDDBCCBCCAABDBAAD
12 | ACBCACCCDADDBADACCADCBCBCBBBDBBABAACCCAACCACBBBDACAAABBCAACA
13 | ABCBADCCACDABCBDDDCDBDABBCBDBDCDDBCDAAABCADBBB AAACBDDDBACADD
14 | BACBCCCDCAAB ACAABACBADCBABBBBCBDBCBBCDBCBACAABDBCADDBBBCACC
15 | DCBBCCACBADDACAABCCADDBBBCABBADBDCBDDBBDBCCDDCBADDCDCDDDDDD
16 | DCACDCCBBDDBCBADDBDCDADABBCBADDDAACBAABCBADBADCACACBDCBCCAAC
17 | CBCADBBCCACCDAAADCCBDBBBBCDAACCCBACDACCDBAADACABBADBDAAABCD
18 | CABAACBCACDAACBBDBCBDDABBCBDBDCDACCBBCCBDDDDACABBAADCCDDDDC
19 | CCBACBADABADCBACAACBCDADCBCCDDBCBACCBDDBCCCAABBDDBAACDBBCCB
20 | ADCDBDDBCBBDACCBDBDCDADABBCDADDDAACCAABCCADBADCACABDDCCCAA
21 | CBCADBBCCACCDAAABCBDBBBAACCCACDACCDBDDDAABBADBDAAABCD
22 | CDABACCCACCCABBDDBADABBBBCDACCACDACCDBAADACABBADBDADDDDC
23 | DDCDADBCCACDABCBADCDDBCBABCDDBABACAABABCBBCBADADDBDBCCDD
24 | ADCDBDDBCBBDACCBDBDCDADABBCDADDDAACCAABCCADBADCACABDDCCCAA
25 | ABBACCCBBAABBDDBCCDADABBCDADDDAACBDABCBADBBDCADABBBACBAD
26 | DBACACCCAADCCCBDBDDBDADACBDDADDADBBBAABCBADBADCACACABDBCCDA
27 | AABBCADDABBCAADACBBBCDCABCDDBAACBACBACCABAAACCBADBADCDDBD
28 | DDABCCCDAAADCCAADACBBADDBBADEBBDABCCABDBADBCAACBDBAADDBACAD
29 | ACCCADCDADBDCCCADACCADADBCDECCADABACBABABCBBCBADACDBDBCCBD
30 | DCCDADADCDCCCADACCACABBCBDBCCDABACBCAABCBBCBAAADDBDBCCDD
31 | DABCBBCDDACDAACAADACBAADBBBCBBBCAAACABCDCCDCCADACADDDDBDBD
32 | CCBBCBDDACDAACAAAACBAADBACBBBCAAACABDDCDDCCADADACADDDDBDDC
33 | DCBAABCDACDAACAADDBAABDBCBBCCAAACABCCDDCCCADADACADDDDBDD
34 | CDBBCCCCACBDACBDDDACDABBCBDBCDAAACCBCCDBCAADDBBDADDDBADDDC
35 | ABCCACCBBAABBDDBCCBADABBCDADDDAACADABCCBDBBDCADABBBACBAD
36 | BDABADBDDBCBACCCBDDADACACBADBDBAACACBAAABBDACCBCCDCBDBCCABC
37 | CDABADBDDBCBACCCBDDADACACBADBDBAACACBAAABBDACCBCCDCBDBCCABC
38 | CDCCCDCAAABACCCDCDCBADABCCDADCAABACBAAABABAACACDDCDDCBCCBAA
39 | CBDCCCCACCCADCBDDCCDADABBCBDBCAADCBBCBCCDBBDDACCACBBBDBABC
40 | ABBAADCBACDAACBBDBCBDBABACBDDCCACDACCBCDDDAABBAADCCDDDDC
41 | CDDAADBCDCDBAABBADACAABBBADAACACACCBCCDBCAAAAADACCCDDDDC
42 | ABDAADCBACDAACABDBCBDBABACDADCCCACCBADCCDCAAAAABBBDBADDDDDC
43 | DCCCCCDAAADCACABBBCCADDBBADCBBCBDBCBCCBDBABABADCAADDCADCA
44 | CDCCDBCBAABACCCDBCCDADABBCDADCAABACBAAABBDAAACBDDCDDCBCCBAA

```


45 | DCCDABCDACBDCCAADBCDBADBDADADCAACDCCBADCBDBAABACCAACBBAADD
46 | CDCADDABDBCBDADBDDBDACCDBBBAAAADADBADCAABDABADEBBAADBDABBBBCDC
47 | CCBBCDCCAAADCACAAAACDABBBBCCBBBCBABADDCDBCCACABBBBBDADDBBCACC
48 | ABCBCCCDAAACACCCDCDCDADDBBADACCBACACCCCBDBDCADCBACDCDBACCBC
49 | CDBBDBCCCAACCAAADCDDBBBBBBAAACACCACDACCDBDDDAABAAABDAABBCDC
50 | DCBBCCBDDAADCCCBAADACBCDABBBADBBBCBACCBBBCBCCBABAADBADBDBCCDCD
51 | DABAADABACDAACBBDBCBCDBABBCBDBDCBACDACCDBAADACABBAADCCDDDDC
52 | CCADBBBCAADCAADBDDDACACCBACDACCACDACCDBCAABAAADADDDCCDDDC
53 | ACCAADCCDDCDAAADDBCBDCDCCBACAAACABAAABADCBABBBADDADCBDBBBADAB
54 | ACCAACCCACDAABDDAAABDCBABDBBCBBAAABADCBABDBABDADCADBBBADAB
55 | DCBABBCDDACDAACAADACBAADBBBCBBCCAAACABCCDDCCCADADCADDDBBBDBD
56 | ABDAADCBACDAACBBDBCBCDBBBBCDBACCCACCBADCBCAAAAABBBDBDADDDDC
57 | CDDAAABCDCCDCCDABDCBACBBAACAAACBACBDBADADACBADACCCDDDDDB
58 | DCBABBCDDACDAACAADACBAADBBBCBBCCAAACABCCDDCCCADADCADDDBBBDBD
59 | DABCBBCDDACDAACAADACBACDBBCCBBCCAAACABCDCCDDCCCADADCADDDBBBDBD
60 | CDCCADCBDBCBDACBADDAADBBBCDADCADCADCAABDDADBBAAABDBABBCDDC
61 | CDCADDABDBCBDADBDDBDACCDBBBAAAADADBADCAABABBAADBBADBDABBBBCDC
62 | CDAA ACCAAACCAADCCAABCA BBBAACDACCADDDBBABA AAAAADDADDDAAAAD
63 | CCBBAACCCACBDADCAADBCBADABBBADACCBABCCACAAADDADBBDDAADCBBDDBC
64 | CCBACCCDACBDADCAADBCBAABBCDACCABCCACAAADDADBAADAACADAADDCC
65 | BDADADBCCCCBBADBBCCCDADBDDBDECAACCBDEDDABDABDDABDDDDDDADDAB
66 | BDADACBDCBBAADBBCCCDDBBDBBDBBACDDDBDBBCABDDABDBDAADACDBB
67 | DBADACCCDBCBDCBBBCCCADABBCDACCADABCAABAADBABBABADDDDDADDAB
68 | DBADACCDDBCBDCBBBCCCADABBCDADCAADCBCAADAADBABABADDDDDADDAB

LAMPIRAN 2: DATA RESPONS PESERTA TES KELOMPOK II

[illegible]

41 | DACBACDACDBAAACBACACABACABBCBACDBACBABBDDBBDDABACAC
42 | CABCCCBABDDCDAACDBCCBBBDCBACAABDABBBBCDAABBDADAACBD
43 | BACCDACDCAACDAABCABDDCBBCBDBCBCCACCCADDDBBDDAD
44 | DACBACDACDBAAACBACACBBBCABBCBACDAACCABBDADDDBADCAA
45 | DCBBDDCCCCACDCCBACACBBBAABBCBAABBCCACABAAABCBADABC
46 | BACCDACDADDACDAABCABDDCBBCBDDCCBCCACC ADDDBBDDAD
47 | BACCDACDCCAACDAABCAAADCDBBDDDDCACCBBADDDDBDDBAD
48 | CCADDCCDC CBCADCDACBDBBDCBBCBCBBBCCACAADABDDBADCAD
49 | DADBDAAACDDBACCBACBCBBAA ABAACAABBCACABAAABBBADCCDC
50 | ACABACBACCBDCDCCBACDCBBBCDBBCBCBBBCBACBADABCDBADCAD
51 | DACBACDACDBAAACBACACABBCACBCBACDAACCABB ADDDBADCAA
52 | CACAACDACCBDACBACACBBACCBBBBACDDDDCABBCCAABDBBBCBC
53 | DDDCCDCCDBCDAACBACCCBBBADACBBCBDDCCACCCAADDBBBCAD
54 | DDACCDACDCCDAABACCCBBBDAACCBBDCBCCABBCADDACBBCAA
55 | DDACCDACDCCDACBACCCBBBDBBBBBDCCBDDCABCCADDDBBBCA*
56 | DBADADAABDCCAADCDCBCCBDDCCCBDDABADBAABBBBCDCBCDA
57 | BACCDACDDBADACBAABCABBCBBCBDDCCCCABC AADCBBBDBA
58 | BACCDACDADD DACBAABCABBCBBCBBBCCCCADDCCAADCBBBBA
59 | CAABACDACDBAAACBACACBBBCACBCBACDBACCABBDADDDBADCAA
60 | BACCDACDADDACBAABCABBCBACBADACCCCBACAAABDCCBBBA
61 | AACACDACBBDDACBACACBACCBBBBAADAAACABBDAAABDDCCBC
62 | CAABDCCDACDBAAACBACADBBBCACBCBACDBACCABBDADDDBADCAA
63 | AACADCDACCBCDACBACACABBCABBCBDBBCCAACBCAADCBBDDBD
64 | AACCDACDACCBDAAACABCACBBBACBBCBBADACCAABACAABABDBCBD
65 | DACBDCDACDBBBACBAABCABBCBBCBDDDDACABCAAADDBABCBD
66 | AACACDACBCDDACBBCACAAACCBBBBAACABACABBDAAABDDBCBBC
67 | AACBACDACBBDDACBACACBBACCBBBBADCBACCABBAAABDB BCBC
68 | AACBACACBACBAACBBCCCBBCABC* BADDCCBABCBBADDBABCDA
69 | AACBACADAADCABCDACBCBBBCABBDDBADCCDAACCABAADABADADA
70 | AACBACDACDBADACBACACBBACCBBBBADCBACCABBAAABDBABCDC
71 | AACADCDACDBADADACAACBBBAABCCAADDBCCCCDCADDDDBDAAC

LAMPIRAN 3: SKOR PESERTA TES KELOMPOK I

1	35.00	44	19.00
2	12.00	45	18.00
3	17.00	46	9.00
4	21.00	47	28.00
5	21.00	48	19.00
6	18.00	49	10.00
7	13.00	50	36.00
8	19.00	51	13.00
9	25.00	52	11.00
10	23.00	53	18.00
11	23.00	54	21.00
12	14.00	55	27.00
13	20.00	56	13.00
14	34.00	57	10.00
15	23.00	58	27.00
16	15.00	59	27.00
17	11.00	60	13.00
18	14.00	61	10.00
19	30.00	62	13.00
20	14.00	63	22.00
21	11.00	64	17.00
22	16.00	65	17.00
23	13.00	66	11.00
24	14.00	67	14.00
25	17.00	68	15.00
26	12.00		
27	15.00		
28	30.00		
29	18.00		
30	18.00		
31	27.00		
32	25.00		
33	25.00		
34	16.00		
35	19.00		
36	14.00		
37	14.00		
38	18.00		
39	14.00		
40	11.00		
41	11.00		
42	12.00		
43	34.00		

LAMPIRAN 4: SKOR PESERTA TES KELOMPOK II

1		32.00	44		34.00
2		28.00	45		20.00
3		24.00	46		31.00
4		29.00	47		25.00
5		25.00	48		19.00
6		23.00	49		17.00
7		23.00	50		24.00
8		21.00	51		35.00
9		17.00	52		28.00
10		23.00	53		30.00
11		26.00	54		27.00
12		24.00	55		28.00
13		25.00	56		15.00
14		23.00	57		31.00
15		26.00	58		29.00
16		27.00	59		31.00
17		27.00	60		28.00
18		21.00	61		26.00
19		28.00	62		30.00
20		35.00	63		31.00
21		27.00	64		26.00
22		17.00	65		31.00
23		16.00	66		25.00
24		12.00	67		30.00
25		23.00	68		19.00
26		18.00	69		20.00
27		13.00	70		31.00
28		15.00	71		25.00
29		17.00			
30		20.00			
31		35.00			
32		17.00			
33		19.00			
34		27.00			
35		26.00			
36		24.00			
37		30.00			
38		15.00			
39		21.00			
40		12.00			
41		31.00			
42		15.00			
43		30.00			

LAMPIRAN 5: IDENTITAS PENELITIAN

Rustam, lahir di Kotabumi Lampung Utara 12 September 1965. Gelar sarjana Pendidikan Matematika diraih tahun 1989 di FKIP Universitas Lampung. Pada bulan September 1997 menyelesaikan Magister Pendidikan Jurusan Penelitian dan Evaluasi Pendidikan (PEP), Program Pascasarjana IKIP Yogyakarta. Sejak tahun 1990 menjadi Pegawai Negeri Sipil di lingkungan FKIP Universitas Terbuka dengan NIP. 131925717, dan sejak 1 April 1999 menyanggah Jabatan Fungsional Lektor Madya.

Penelitian yang pernah dilakukan dan terkait dengan penelitian ini adalah:

1. Karakteristik Tes Program Penyetaraan D2 PGSD Universitas Terbuka dengan Model Rasch (1997).
2. Kualitas Tes Program Penyetaraan D2 PGSD Universitas Terbuka (1999)
3. Konsistensi Parameter Tes (1999).

Artikel yang pernah dipublikasikan dan terkait dengan penelitian ini adalah:

1. Karakteristik Tes Program Penyetaraan D2 PGSD Universitas Terbuka: Implementasi Model Rasch (1998). Jurnal PPs IKIP Yogyakarta.
2. Konsistensi Parameter Tes (dalam proses). Jurnal Pendidikan Universitas Terbuka.